Illumination invariant stationary object detection

Waqas Hassan, Philip Birch, Bhargav Mitra, Nagachetan Bangalore, Rupert Young, Chris Chatwin

School of Engineering and Informatics,
University of Sussex, Falmer, Brighton, BN1 9QT, United Kingdom
Phone: 44 1273 877207
w.hassan@sussex.ac.uk, p.m.birch@sussex.ac.uk, b.k.mitra@sussex.ac.uk, nb56@sussex.ac.uk, r.c.d.young@sussec.ac.uk, c.r.chatwin@sussex.ac.uk

Abstract

A real-time system for the detection and tracking of moving objects that becomes stationary in a restricted zone. A new pixel classification method based on the Segmentation History Image (SHI) is used to identify stationary objects in the scene. These objects are then tracked using a novel adaptive edge orientation based tracking method. Experimental results have shown that the tracking technique gives more than a 95% detection success rate, even if objects are partially occluded. The tracking results, together with the historic edge maps, are analyzed to remove objects that are no longer stationary or are falsely identified as foreground regions due to sudden changes in the illumination conditions. The technique has been tested on over seven hours of video recorded at different locations and time of day, both outdoors and indoors. The results obtained are compared with other available state of the art methods.

1. Introduction

Identification and tracking of moving objects that become stationary is an important part of many video surveillance applications. Two such applications are the detection of illegally parked vehicles and abandoned object detection. A number of techniques have been presented in the literature in the recent past to identify and track stationary objects [1-4,6-23]. Although partially successful many of the techniques fail on one or more of the following scenarios: the image is too crowded, the ambient lighting changes, they are too slow to compute, static objects become part of the background model.

The core of the problem is to deal with objects that are stationary for a long duration of time. This is especially challenging in situations such as parked vehicle detection where there is changes in illumination conditions that changes the appearance of the object over time. In addition, the object is likely to become occluded by other objects such as passing vehicles or people.

A solution based on these challenges is proposed in this paper. It consists of two stages: the segmentation of the new static objects and then the tracking of the stationary objects over a number of video frames. The proposed method not only tackles lighting issues in variable outdoor scenes but also handles occlusions. It has been demonstrated with the parked vehicle and abandoned baggage detection but it could be adapted to other scenarios.

2. Stationary Object Detection

The first stage of the proposed method is to accurately identify new stationary objects. We make several assumptions about the data: the stationary objects are initially not part of the scene and come in at a later stage; they must be static in the scene for at least 10 seconds before being labeled as stationary and, finally, they must fall within an user defined zone. To identify objects that are newly stationary in the scene we must first locate moving objects and then trigger the tracking stage when these have become stationary. Stationary objects detection is therefore a two-step process:

- An initial segmentation of the video sequence using a Gaussian Mixture Model (GMM) to determine moving objects;
- A segmentation history image (SHI) is created using this output of the GMM to filter out transitory motion and determine when an object has become stationary;

For outdoor systems changes in the illumination conditions or the starting and stopping of vehicles in busy traffic scenes will cause difficulties in segmentation. To cope with this GMM background models must be updated. If the model is updated at a low rate any sudden change in illumination can result in the identification of false foreground areas. If the update rate is higher to cope with the illumination changes, stationary objects become part of the background model very quickly resulting missing segmentation pixels.

Partial or complete occlusion is also a challenge. Most solutions proposed in the literature require an isolated view of the object before identifying it as abandoned. This requirement is typically difficult to meet, especially in crowded scenes where abandoned objects are frequently occluded.

In our system we have set the update rate to be high to cope with rapid lighting variations but we have also introduced an SHI to post-filter the GMM output and determine when an object has become stationary.

The output of the GMM produces a set, Ω^t, of foreground pixels at time t. The SHI data, Γ, is initially zero then increased by one for every foreground pixel per frame:

$$\Gamma_{\Omega}^{t} = \Gamma_{\Omega}^{t-1} + 1 \tag{1}$$

Next the objects in motion are removed whilst keeping stationary pixels. This is achieved by setting the SHI to zero for every pixel not in the current Ω^t:

$$\Gamma_{\overline{\Omega}}^{t} = 0 \tag{2}$$

where the bar indicates all pixels not in the current set Ω. This eliminates all those regions that are moving from the SHI, thus leaving only those pixels that belong to the stationary objects.

Finally, a pixel is declared as potentially stationary when its SHI value is greater than some threshold,Λ. Λ is simply the number of frames we wish to count before declaring a pixel stationary. In our experiments, we have set this equal to 10 seconds of video. The stationary pixels are described by the set *S*:

$$S^t = \{a : \Gamma_a > \Lambda\} \tag{3}$$

where *a* is all the possible coordinates in the image. Because we are selecting static objects with a pixel level SHI, our solution does not require the object to be completely isolated before it determined to be a new static object.

The SHI effectively acts as a counter to determine when an object has become stationary, but in addition, by the careful selection of the GMM update rate it can filter out false objects due to lighting variations. It must be set so that the object remains the foreground for at least Λ frames. It is noted that intensity changes from lighting variations are typically much smaller than the intensity changes from vehicles parking over a short time frame. Thus, the update rate of the GMM can be set such that lighting variations have become part of the background and are thus removed from the SHI (via equation 2), before the Λ frames have been reached. The choice of the update rate was selected manually but it is robust for many different videos and only need changing if Λ is altered. It does not need changing to cope different weather conditions or from switching from day to nightime lighting. This has been tested over 24 hours of video sequences from multiple sources and has shown that in over 95% of the test sequences the effects of sudden illumination change can be removed this way.

After the pixel segmentation, a connected component algorithm is run on the binary image, S^t, to produce objects, Θ_n^t where n is the connected component index and $n \in \{1, \ldots, N\}$ where N is the number of connected components discovered. Objects that are too small or too large are then removed.

3. Stationary Object Tracking

A new adaptive edge orientation based technique is proposed in this section to track objects that are identified as stationary. To maintain the tracking, each object at time t is compared against the object at time T, where T is the time the object is first flagged as stationary by the SHI (equation 3). A correlation based method has been developed that uses the direction of the edges to indicate the strength of the match rather than the edge strength alone. A 2-D spatial gradient measure is performed on an image flagged as a stationary object, $I_{\Theta_n}^T$, where I is the monochrome image intensity. This is performed by a 3x3 Sobel convolution operator and is used to find the edges in both the vertical and the horizontal directions. So that the edge directions are well defined, the magnitude of the gradient is then calculated at T. This is done by taking the square root of the sum of the squares of each direction gradient result in Θ_n^T. Finally the output is thresholded to form a subset of coordinates, M_n. The threshold is chosen such that most of the edges are preserved, ie only pixels with very small gradients and hence poorly defined directions are removed. This value, along with the other thresholds does not need changing between dataset. The set M_n does not change until the object moves away from the scene and it is removed from consideration, i.e. all future comparisons will be made against this same set of pixels.

The direction of the gradient of each member of M_n is then calculated by taking the arctangent of the ratio the vertical and horizontal gradients. We denote this as $\Phi^T_{M_n}$, where $\Phi^T_{M_n}$ can range between 0 to 360 degree.

The obtained angles are then used to calculate a metric that gives a measure of the degree of correlation between an object at the current time *t* and the historic time *T*. For the n^{th} object this is the real component of:

$$C^t_n = \sum_{M_n} \frac{\exp\left(i\Phi^t_{M_n}\right)\exp\left(-i\Phi^T_{M_n}\right)}{\left|M^T_n\right|} \tag{4}$$

where the vertical bars represent the cardinality of the set, and *T* is the time at which the reference edge mask is produced. We note here that since equation 4 is effectively a high pass filtered correlator since it is matching edges it is very sensitive to movement of the object and exhibits no shift invariance. Any slow moving objects that have been incorrectly labeled as stationary by the SHI process will therefore not correlate and will be rejected. One problem that does arise is when the tracked object is partially or completely occluded. The occlusion problem is discussed in more detail in the following section.

3.1 Tracking under occlusion

The proposed technique overcomes the problem of occlusions by assuming that any current foreground segmented object (from the GMM in stage 1) is a new occluding object and it should be removed from the set of valid pixels used to calculate C_n^t. The occluded region is removed from the mask, M_n, by subtracting the current GMM segmentation image, Ω_n^t. The adaptive mask for the n^{th} object is a subset of the original mask:

$$A_n^t = M_n - \Omega_n^t \tag{5}$$

And the improved correlation result is then:

$$\hat{C}_n^t = \sum_{A_n} \frac{\exp\left(i\Phi_{A_n}^t\right)\exp\left(-i\Phi_{A_n}^T\right)}{\left|A_n^t\right|} \tag{6}$$

The object is assumed to be tracked when the $\hat{C}_n^t$ is greater than 80% of the auto-correlation value. The 80% was selected as balance between rejecting image noise and avoiding the tracking false objects, but it appears to be reasonably insensitive to changes in its value. The process works efficiently of situations where there is partial occlusion. But if object is completely occluded during the tracking process, the proposed method waits for a user defined number seconds before dropping the object. For situations where a stationary object is occluded by another object that become stationary, the algorithm will drop the previously occluded object and will start tracking the new object.

Fig. 1 shows images highlighting the effects of occlusion and how the proposed technique is used to obtain an adaptive edge map and binary mask for the occluded tracked objects.

The proposed method was tested on the data presented in Fig. 2. The results obtained are compared with the image intensity based cross-correlation technique presented in [3]. Table 1 shows that the adaptive edge orientation based technique gives over 95% matching results even if the object is partially occluded as compared to 69% matching results obtained using the image intensity based cross-correlation technique, hence demonstrating that proposed method is more reliable for tracking objects under occlusion.

The authors note that if the same adaptive approach was to be applied to other cross-correlation based techniques there should also be an improvement in performance over that which has been reported to date.

3.2 Tracking under changing lighting conditions

The ability to track objects in changing illumination conditions is another significant characteristic of the proposed method. For surveillance systems deployed in outdoor environments, it is very important to accommodate situations where lighting is not consistent. A change in lighting can not only cause shadows to appear that can falsely classify pixels as members of $\mathbf{\Omega}$, but it can also change the apparent color of a tracked object. This can lead to false tracking results if objects are tracked based on color features.

Techniques based on the energy or magnitude of the edges, e.g. in [5], can also suffer from changes in illumination. A change in lighting can cause the magnitude of an edge to change which can result in false tracking outputs. Since our adaptive edge orientation based technique considers the orientation of the edge rather than the intensity, and there is no dependency on color features, it performs better in conditions where lighting is not consistent, as we shall now show. The results below show that edge direction is more stable under lighting changes than both edge magnitude and colour. The proposed solution was applied to the highly variable lighting video sequence shown in Fig. 3. The results are compared with two other tracking techniques based on color in [6] and edge energy in [5]. The running average matching score obtained over 800 frames are presented in Fig. 4. This running average matching score is obtained by adding $\widehat{C}_n^t$ for proposed method, the cross-correlation match [5] and the I-MCSHR similarity [6], separately for each frame before averaging it out over total number of frames the object is stationary. It can be seen from Table 2 that the matching results obtained from our proposed solution are consistently better than the other two techniques throughout the sequence, especially when there is a large change in the illumination conditions.

4. Object Removal

Our algorithm deals with the removal of objects that are no longer stationary and with those regions that are falsely identified as static objects. Historic edge maps and tracking results are analyzed for this purpose. Edge maps for the alarm zones are obtained using the technique proposed in Section 3 and stored historically. The frequency of this storage process depends on the storage capacity of the system implemented; it has been kept to 15 minutes worth of video sequence in our case. For any foreground regions that are labeled as stationary, edge maps are calculated and matched with the historic edge map database for the exact pixel location. The idea is that for any legitimate object, the obtained edge map will always be different from edge maps stored in the database for the exact pixel location. Since the edge directions are robust against lighting changes the storage of the maps for comparison has proved to be very robust. Experimental results have shown that although the technique makes the process more computationally expensive, it significantly reduces the number of false alarms.

The system might fail if a similar shape and sized object is parked at the same location at a different time, but this is unlikely since equation 6 is not shift invariant, ie, the new object has to be parked at exactly same pixel location and it has to be identical. In addition it has to be parked within 15 minutes of the first object left the scene. This is because the historic edge database is progressively refreshed such that the refresh process completes after every 15 minutes

After the above matching process, regions identified as legitimate foreground objects are tracked using the adaptive edge orientation based technique and the results are compared against a chosen threshold. It is observed that when a stationary object pixels move, the orientation of the edge map rapidly changes and causes the matching results to drop below the selected threshold. This can result of two different scenarios: the tracked object has either left its stationary position or is occluded by another object. To distinguish between these two situations, the binary mask overlapping the object's stationary position is obtained from most recent GMM output and analyzed. If the mask is originated from with in the area of Θ_n^T in terms of pixel location (ghost object) and its total area is also within the area of the stationary object, the object is assumed to have moved and is no longer tracked. On the other hand if that is not the case, the object is considered to be occluded and the adaptive edge orientation technique is applied to keep track of the stationary object.

5. Experimental Results

The proposed technique has been tested on two different datasets divided into two scenarios: illegally parked vehicle detection in outdoor scenes and abandoned object detection at train stations. In both cases the alarm events are generated after 60 seconds of object being left stationary. The update rate was selected such that legitimate foreign objects were categorized as foreground for at least Λ frames before becoming part of the background. This dependency on Λ frames helps our proposed system not only pick up stationary objects efficiently but also reduces any affect of sudden illumination change.

An 80% matching threshold was selected in order to track the object using the adaptive edge orientation technique. This value appeared to work well across all the tested environments with changing lighting conditions, and for any video rate and image quality. If the threshold was set too high or too low, the tracking results are compromised on platforms with lower image resolution or poor video quality (some sequences from *i-LIDS* dataset). All thresholds were kept fixed through out the testing of the proposed system across all video sequences with different video resolution and quality, viewing angles, lighting conditions (day or night) and scene density. The algorithm has been implemented in C and runs at over 15 frames per second on a single core 2.66 GHz PC. The results obtained are also compared with other state of the art methods presented in the literature.

5.1 *i-LIDS* Dataset

The algorithm was evaluated on two sets of i-LIDS [24] *Parked Vehicle* video sequences. Both sets contain three progressively more demanding video sequences, taken at different times of the day. The first set consists of long video sequences with multiple alarm events; the second set contains short video sequences of 2-3 minutes duration with each containing a single alarm event. In both sets the alarm events are generated within 10 seconds after a vehicle being parked in a prohibited parking zone for 60 seconds.

Out of 105 alarm events, the method successfully detected 99 events so an overall detection rate of 94% was achieved by our algorithm. However, there were 18 false alarms i.e. 15% of the events. This was due to the fact that no explicit vehicle detection technique has been used to identify whether the detected object is a vehicle or not. Fig. 5 shows images from the video sequences tested. The algorithm has been comprehensively tested in different lighting conditions, various street views and with varying video quality without changing any threshold values.

Table 3 shows the short videos [25] results in comparison to other authors' results. The average error is calculated by comparing the ground truth start and end times with the acquired results. The system detects illegally parked vehicles more accurately than any other available method.

The proposed method was also tested on three *Abandoned Baggage* video sequences from the *i-LIDS* dataset [25]. We focused on alarm generation if the object is left stationary for over 60 seconds irrespective of whether it had been left unattended or not. All the events were successfully detected by our system with a low number of false positives. The two false alarms identified in the "hard video" sequence are, in fact, bags on the floor with their owners sitting next to them. Since we are not associating people with bags, we put them in the false positive category.

5.4 *PETS* 2006 Dataset

Fig. 6 shows images from the *PETS 2006* dataset [26] in which the abandoned object is detected. The overall results of our proposed algorithm when tested on these sequences are summarized in Table 4 in which comparisons are also made with the other available state of the art methods. Table 4 shows that the proposed method picks up all the genuine alarms without generating any false alarms. When compared with the other available techniques, it can be seen that the method produces very accurate results (equal to those reported by [22]). Although [23] also produced 100% results no information about false alarms was reported by the authors. Although not included in the PETS datasets, the algorithm should cope well with lights switching on and off since it copes with large changes of outdoor lighting.

6. Conclusion

A robust stationary object detection system has been presented in this paper. Two major concerns faced by such techniques, namely lighting variations and occlusion, were addressed in this study. A new pixel classification method based on GMM is initially used to detect stationary objects where SHI was utilized to overcome difficulties caused by variable lighting effects. Once objects are identified as stationary, an adaptive edge orientation based tracking technique is applied to track objects under severe lighting conditions and occlusion. The study has demonstrated that the proposed tracking method has an accuracy of over 95% matching results under occlusion as compared to less than 70% accuracy achieved by the other available state of the art methods. Finally, objects that are no longer stationary or are falsely identified as stationary objects are removed by analyzing historic edge maps and tracking results.

Over 7 hours of video have been tested with 95% of the 121 genuine alarms detected in demanding lighting conditions at a rate of over 15 frames per second.

References

[1] Fatih Porikli, Yuri Ivanov,Tetsuji Haga: Robust Abandoned Object Detection Using Dual Foregrounds. EURASIP Journal on Advances in Signal Processing, vol. 2008, 1-10 (2008)

[2] Lee, J.T., Ryoo, M.S. , Riley, M., Aggarwal, J.K..: Real-Time Illegal Parking Detection in Outdoor Environments Using 1-D Transformation. IEEE Transactions on Circuits and Systems for Video Technology, vol.19, no.7, 1014-1024 (2009)

[3] Tian, Y., Feris, R., Liu, H., Hampapur, A., Sun, M.: Robust Detection of Abandoned and Removed Objects in Complex Surveillance Videos. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews , vol.PP, no.99, 1-12 (2010)

[4] Venetianer, P.L., Zhang, Z., Yin W., Lipton, A.J.: Stationary target detection using the objectvideo surveillance system. IEEE Conference on Advanced Video and Signal Based Surveillance, 242-247 (2007)

[5] Hassan, W., Mitra, B., Chatwin, C., Young, R., Birch, P.: Illumination invariant method to detect and track left luggage in public areas. SPIE 7696, vol. 76961 (2010)

[6] Piccardi, M., Cheng, E.D.: Multi-Frame Moving Object Track Matching Based on an Incremental Major Color Spectrum Histogram Matching Algorithm. IEEE

Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 19-19 (2005)

[7] Bevilacqua, A., Vaccari, S.: Real time detection of stopped vehicles in traffic scenes. IEEE Conference on Advanced Video and Signal Based Surveillance, 266-270 (2007)

[8] Boragno, S., Boghossian, B., Black, J., Makris D., Velastin, S.: A DSP-based system for the detection of vehicles parked in prohibited areas. IEEE Conference on Advanced Video and Signal Based Surveillance, 260-265 (2007)

[9] Guler, S., Silverstein, J.A., Pushee, I.H.: Stationary objects in multiple object tracking. IEEE Conference on Advanced Video and Signal Based Surveillance, 248-253 (2007)

[10] Lee, J.T., Ryoo, M.S., Riley, M., Aggarwal, J.K.: Real-time detection of illegally parked vehicles using 1-D transformation. IEEE Conference on Advanced Video and Signal Based Surveillance, 254-259 (2007)

[11] Tian, Y., Feris, R.S., Hampapur, A.: Real-time detection of abandoned and removed objects in complex environments. IEEE Int. Workshop Visual Surveillance (2008)

[12] Singh, V.K., Atrey, P.K., Kankanhalli, M.S.: Coopetitive multi-camera surveillance using model predictive control. Journal of Machine Vision and Applications, vol. 19, no. 5–6, 375–393 (2008)

[13] Arsi´c, D., Hofmann, M., Schuller, B., Rigoll, G.: Multi-camera person tracking and left luggage detection applying homographic transformation. Tenth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance,

PETS 2007 (2007)

[14] Auvinet, E., Grossmann, E., Rougier, C., Dahmane, M., Meunier, J.: Left-luggage detection using homographies and simple heuristics. IEEE International Workshop on *PETS*, 2006, 51–58 (2006)

[15] Krahnstoever, N., Tu, P., Sebastian, T., Perera, A., Collins, R.: Multi-view detection and tracking of travelers and luggage in mass transit environments. IEEE International Workshop on *PETS*, 2006, 67–74 (2006)

[16] Ferrando, S., Gera, G., Massa, M., Regazzoni, C.: A New Method for Real Time Abandoned Object Detection and Owner Tracking. IEEE International Conference on Image Processing, 3329-3332 (2006)

[17] Tang Yi Ping, Feihu Hu, Guanbao Yang, Bei Lin: Property protection system based on Omni-Directional Vision Sensors. 3rd International Congress on Image and Signal Processing (CISP), 181-184 (2010)

[18] Jianting Wen, Haifeng Gong, Xia Zhang, Wenze Hu: Generative model for abandoned object detection. 16th IEEE International Conference on Image Processing (ICIP), 853-856 (2009)

[19] Xuli Li, Chao Zhang, Duo Zhang: Abandoned Objects Detection Using Double Illumination Invariant Foreground Masks. 20th International Conference on Pattern Recognition (ICPR), 436-439 (2010)

[20] Qiujie Li, Yaobin Mao, Zhiquan Wang, Wenbo Xiang: Robust Real-Time Detection of Abandoned and Removed Objects. Fifth International Conference on Image and Graphics, 156-161 (2009)

[21] Tang Yi Ping, Chen Yao Yu: Illegally Parked Vehicles Detection Based on Omnidirectional Computer Vision. 2nd International Congress on Image and Signal Processing, 1-5 (2009)

[22] Jing-Ying Chang, Huei-Hung Liao, Liang-Gee Chen: Localized Detection of Abandoned Luggage. EURASIP Journal on Advances in Signal Processing, vol. 2010, Article ID 675784, 9 pages (2010)

[23] Lv F., Song X., Wu B., Singh V.K., Nevatia, R.: Left luggage detection using bayesian inference. 9th IEEE International Workshop on Performance Evaluation in Tracking and Surveillance (*PETS* '06), 83-90 (2006)

[24] *i-LIDS* dataset, United Kingdom Home Office dataset. Available at http://tna.europarchive.org/20100413151426/scienceandresearch.homeoffice.gov.uk/hosdb/cctv-imaging-technology/*i-LIDS*/index.htm

[25] *i-LIDS* dataset for AVSS 2007,. Available at ftp://motinas.elec.qmul.ac.uk/pub/iLids/

[26] *PETS* 2006 dataset, Available at http://www.cvg.cs.reading.ac.uk/*PETS*2006/data.html

Figure Captions

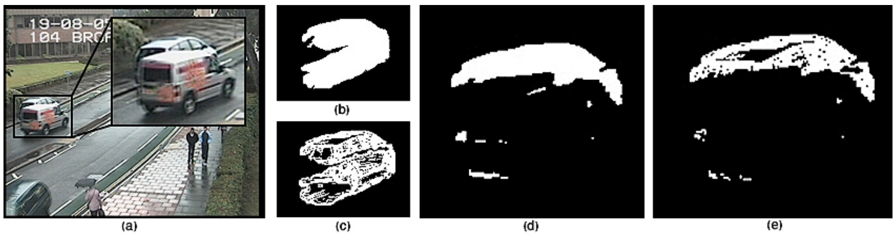

Fig. 1 Effects of occlusion in a typical scene from the *PETS* 2006 dataset; (a) typical scene (b) original object mask, Ω^T (c) object edge mask, $\Phi^T_{M_n}$ (d) adaptive binary mask, A^t_n (e) adaptive edge mask after applying the proposed technique, $\Phi^t_{A_n}$.

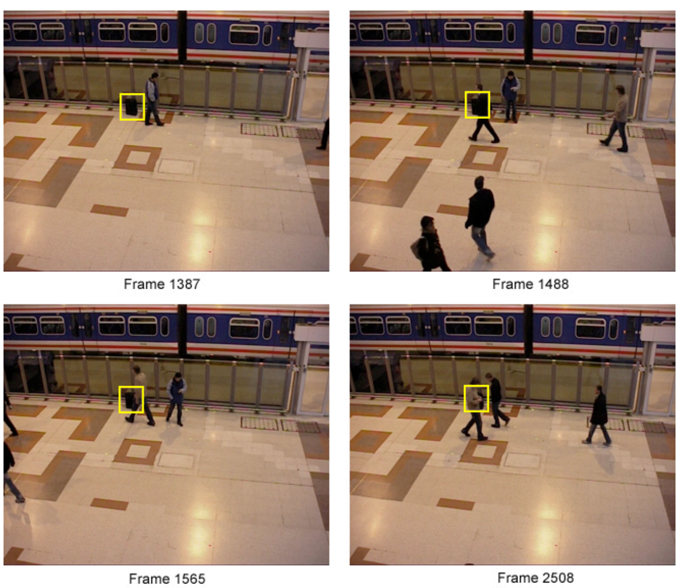

Fig. 2 Video sequence from the *PETS* 2006 dataset used to compare results from the adaptive edge orientation based technique with cross-correlation as used in Tian *et al.* [3]. The tracked object is occluded in Frame 1488, 1565 and 2508.

Fig. 3 Data sequence showing change in illumination conditions. The tracked object is marked in Frame 900.

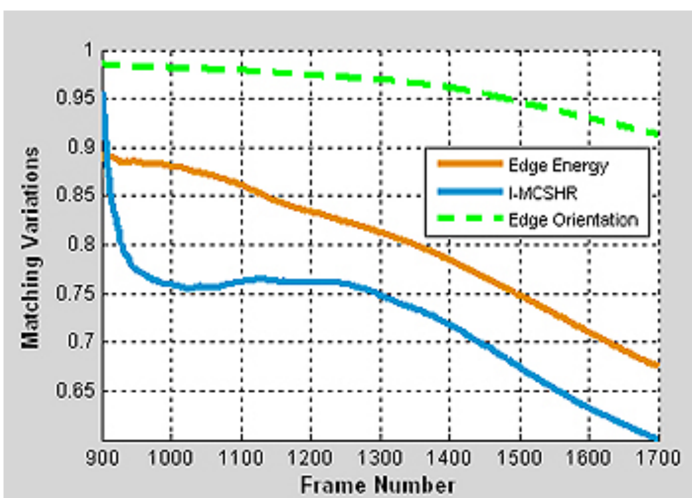

Fig. 4 Running average variation in matching from frame 900 to frame 1700 for the three tracking techniques indicated in the graph. Matching Variation = 1 corresponds to an exact match.

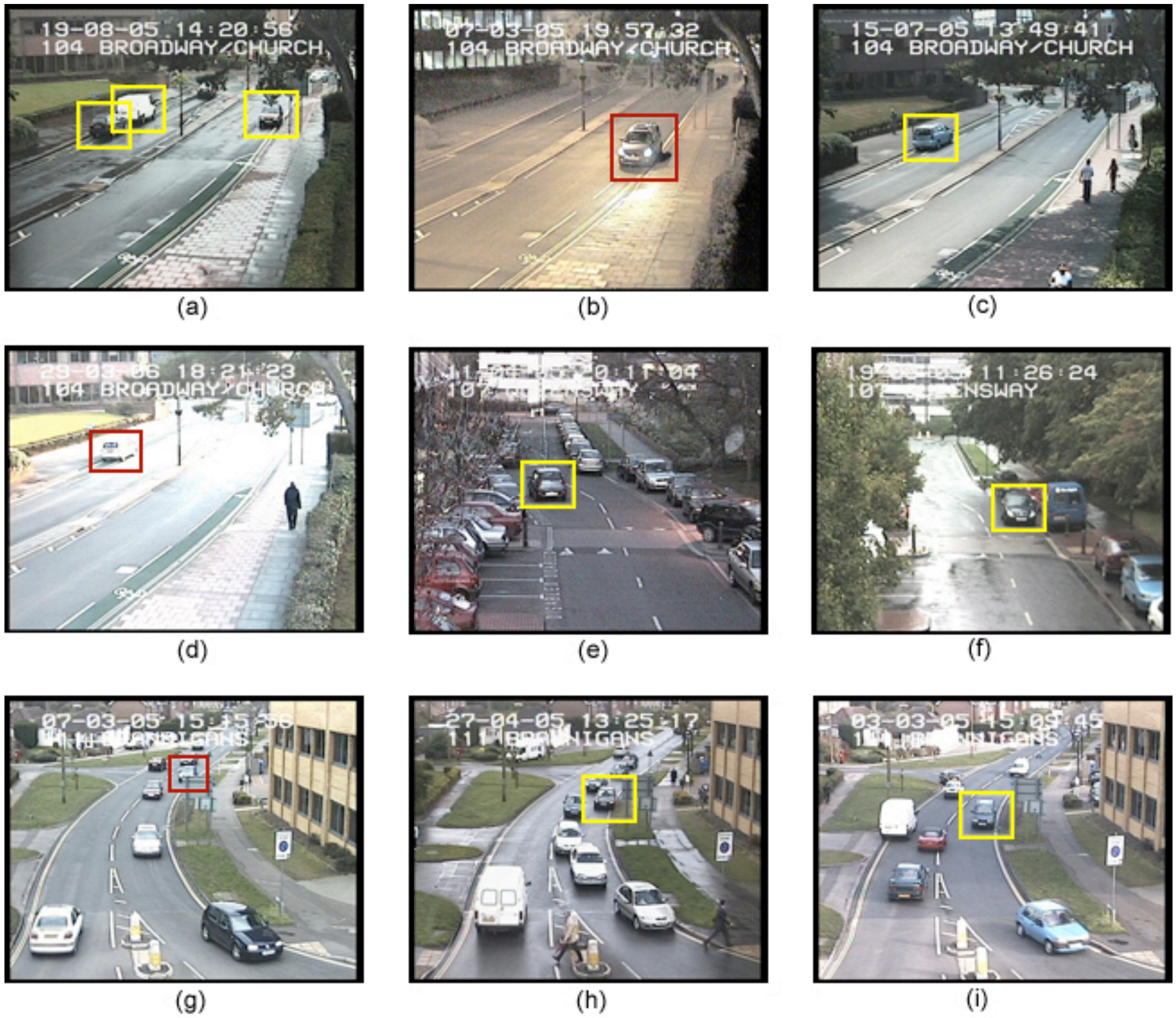

Fig. 5 Images from the video sequences from the *i-LIDS* dataset. The video sequences are recorded at different time of the day with different weather conditions.

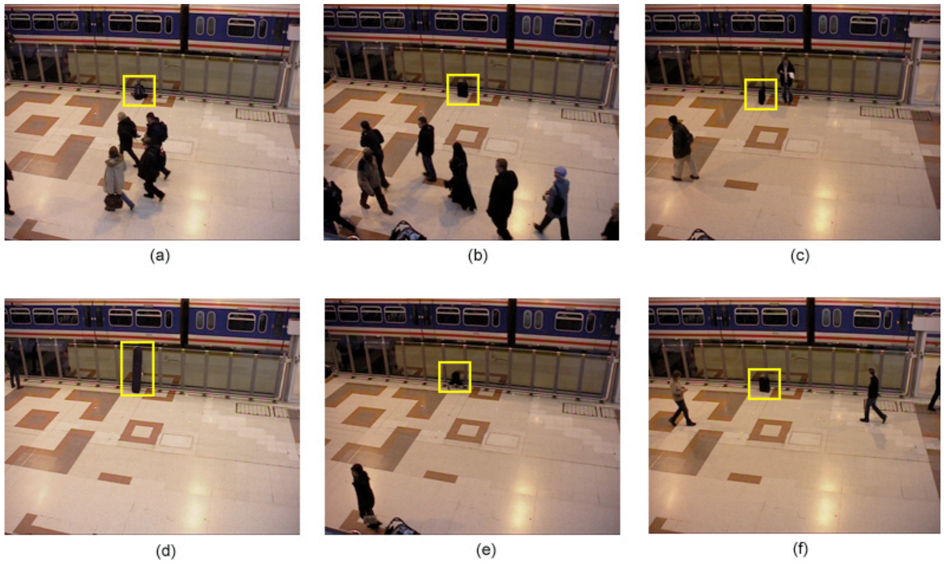

Fig. 6 Images from the *PETS 2006 Dataset* (a) S1 (b) S2 (c) S4 (d) S5 (e) S6 (f) S7. All events are detected successfully with no false positives.

Tables

Table 1: Matching results after applying our adaptive edge orientation based technique and the cross-correlation technique to the sequence shown in Fig. 1

Frame Number	**Adaptive edge orientation match**	**Cross-correlation match [3]**
1387	99.92%	99.83%
1488	96.59%	85.37%
1565	98.36%	94.01%
2508	99.70%	69.04%

Table 2: Running average matching results after applying the adaptive edge orientation based technique; I-MCHSR and edge energy based tracking techniques.

Frame Number	**Adaptive edge orientation tracking**	**I-MCHSR [6]**	**Edge energy tracking [5]**
900	98.53%	88.54%	95.68%
1200	97.40%	83.43%	76.06%
1400	96.08%	78.51%	71.89%
1700	91.30%	67.54%	60.09%

Table 3: Results of the proposed method, along with other techniques, applied to the *i-Lids* short videos. *: sequence is not considered for average error calculation. N/A: data was not available from the author of the technique.

Method	Sequence	Ground Truth		Duration (sec)	Obtained Results		Duration (sec)	Error (sec)	Average Error (sec)
		Start Time	End Time		Start Time	End Time			
Proposed Method	Easy	02:48	03:15	00:27	02:50	03:17	00:27	4	3.75
	Medium	01:28	01:47	00:19	01:28	01:49	00:21	2	
	Hard	02:12	02:33	00:21	02:16	02:36	00:20	7	
	Night	03:25	03:40	00:15	03:25	03:38	00:13	2	
Bevilacqua and Vaccarri [7]	Easy	02:48	03:15	00:27	N/A	N/A	00:31	4	4.33
	Medium	01:28	01:47	00:19	N/A	N/A	00:24	5	
	Hard	02:12	02:33	00:21	N/A	N/A	00:25	4	
	Night*	03:25	03:40	00:15	N/A	N/A	N/A	-	
Boragno *et al.* [8]	Easy	02:48	03:15	00:27	02:48	03:19	00:31	4	5.25
	Medium	01:28	01:47	00:19	01:28	01:55	00:27	8	
	Hard	02:12	02:33	00:21	02:12	02:36	00:24	3	
	Night	03:25	03:40	00:15	03:27	03:46	00:19	6	
Lee *et at.* [2]	Easy	02:48	03:15	00:27	02:51	03:18	00:27	6	6.25
	Medium	01:28	01:47	00:19	01:33	01:52	00:19	10	
	Hard	02:12	02:33	00:21	02:16	02:34	00:18	5	
	Night	03:25	03:40	00:15	03:25	03:36	00:11	4	
Guler *et al.* [9]	Easy	02:48	03:15	00:27	02:46	03:18	00:32	5	6.75
	Medium	01:28	01:47	00:19	01:28	01:54	00:26	7	
	Hard	02:12	02:33	00:21	02:13	02:36	00:23	4	
	Night	03:25	03:40	00:15	03:28	03:48	00:20	11	
Venetianer *et al.* [4]	Easy	02:48	03:15	00:27	02:52	03:16	00:24	5	9.33
	Medium	01:28	01:47	00:19	01:43	01:47	00:04	15	
	Hard	02:12	02:33	00:21	02:19	02:34	00:15	8	
	Night*	03:25	03:40	00:15	03:34	N/A	N/A	-	
F Porikli [1]	Easy*	02:48	03:15	00:27	N/A	N/A	N/A	-	11
	Medium	01:28	01:47	00:19	01:39	01:47	00:08	11	
	Hard*	02:12	02:33	00:21	N/A	N/A	N/A	-	
	Night*	03:25	03:40	00:15	N/A	N/A	N/A	-	
Lee *et at.* [10]	Easy	02:48	03:15	00:27	02:52	03:19	00:27	8	12.33
	Medium	01:28	01:47	00:19	01:41	01:55	00:14	21	
	Hard	02:12	02:33	00:21	02:08	02:37	00:29	8	
	Night*	03:25	03:40	00:15	N/A	N/A	N/A	-	

Table 4: Results obtained after testing the proposed method along with other published methods on all the sequences from the
PETS 2006 dataset. N/A: data was not available from the author of the technique.

Methods	**Total Alarms**	**True Positive**	**False Positive**	**False Negative**
Proposed Method	6	6	0	0
Tian *et al.* [3]	6	6	1	1
Auvinet *et al.* [14]	6	6	5	0
Chang *et al.* [22]	6	6	0	0
Lv *et al.* [23]	6	6	N/A	0